A Guide to Internet Web Sites of CI Interest

This Guide to Internet Web Sites of Counterintelligence Interest is, by no means, meant to be definitive. New Web sites are created daily, making a definitive listing virtually impossible. In fact, this guide is meant to serve as nothing more than a basic 'bibliography' of unclassified web sites relating to counterintelligence issues. Through exploration of the Web sites provided and imaginative searches on 'search engines,' the reader's personal listing of useful and related web sites will rapidly expand far beyond what is given here.

Readers are encouraged to share their Internet Web site discoveries with the NCIX for inclusion in future updates of the guide.

THE INTERNET AND THE WORLD WIDE WEB (WWW)

The Internet is a loose association of thousands of networks and millions of computers across the world that all work together to share information. The World Wide Web is a subset of the Net, i.e. a collection of interlinked documents that work together using a specific Internet protocol called HTTP (hypertext transfer protocol). In other words, the Net exists independent of the Web, but the Web can't exist without the Net.

No one person or organization can claim sole credit for the Internet. However, in 1967 the US DOD Advanced Research Projects Agency (ARPA) designed what became known as the ARPANet computer network as the first design for a worldwide network. In 1990 the US Government officially decommissioned ARPANet, and the National Science Foundation (NSF) took over the role of managing the Internet backbone, which was then called the NSFNet. In 1995, the NSF turned the Internet backbone over to a consortium of commercial providers.

No one person, company, institution, or government organization owns the Internet. No one source foots the bill, governs it, or even has a controlling interest. The Internet is truly a collaborative, collective enterprise.

The Web began in 1989 as a means of improving communications between research scientists around the world. The defining feature of the Web is its ability to connect pages to one another with hyperlinks. Just click on a link, and suddenly you're at a Web site on the other side of the world. (Before the Web, you had to type in exact Net addresses or go through a series of menus to get where you wanted to go.)

As of 1 December 1999, the World Wide Web contains an estimated 800 million Web pages, and the Web is expected to grow by the power of 10 over the next two to three years!

Just a few years ago, the common perception was that computers were the domains of nerds, geeks and society's marginalized individuals. Today, computer acceptance is such that 50% of American homes have at least one computer, and virtually all of them can access the World Wide Web. Retirement homes are encouraging use of computers and the Internet, as a means of preventing the elderly from becoming isolated from society.

Private individuals are now using computers routinely for personal communications, and shopping; they are now their own stock market brokers, bankers, and travel agents. The World Wide Web is making newspapers and magazines obsolete. The 1998 Christmas shopping season netted approximately $3 billion in on-line sales, and in 1999 during the same six week period, close to $6 billion in on-line sales were registered.

While the threats, challenges and opportunities posed by computers (and the Internet) are still being discovered, the ability of the US Intelligence Community to access a virtually limitless supply of unclassified data and information is available now.

Valuable, open-source counterintelligence (CI) resources are available on the Internet. Governmental sites, private commercial sites, and private individual Web sites are on the World Wide Web (www), which is a user-friendly part of the global network of computers that is part of the Internet. Navigation through the Web is simplified by use of search engines, Web directories, and parallel and metasearch sites. Hundreds, if not thousands, of Web sites currently contain valuable counterintelligence information.

As mentioned above, there are three primary types of search sites on the Web: search engines, Web directories, and parallel and metasearch sites.

Search engines, such as Excite and HotBot, use automated software called 'Web crawlers' or 'spiders.' These programs move from Web site to Web site, logging each site title, URL (Uniform Resource Locator, i.e. address), and at least some of its text content. The result is a long list of Web sites placed in a database, which users search by typing in a keyword or phrase. All the major search engine sites now have built-in topical search indexes.

Web directories, such as Yahoo and Magellan, offer an editorially selected, topically organized list of Web sites. To accomplish that goal, these sites employ editors to find new Web sites and work with programmers to categorize them and build their links into the site's index. Most Web directories have now added a keyword search.

Parallel and metasearch sites ride piggyback on the Web crawler sites. Parallel search programs, such as WebFerret, launch simultaneous searches on all the popular search engine sites, returning all the results in a single window. Metasearch sites , such as Metafind, go a step further, in that they let you enter a term in a single field and then automatically account for all the particulars for half a dozen or more popular search sites.

Just like conducting research in a library, when searching for information on the Internet, there is a simple methodology to follow, in order to best utilize the available resources. In a library, you would utilize the card files first, checking first by subject, then by titles and authors. Once you had this information, you would "go to the stacks" and put hands-on the books you've selected, in order of most promising to least promising. Your 'card files' on the Internet are called 'search engines.'

To refine your Internet research techniques, spend some time at these sites, which are guides and/or tutorials:

1. Help Web: www.imaginarylandscape.com/helpweb/

   A beginners's level guide to getting started on the web.

2. Intro to Search Engines: www.kcpl.lib.mo.us/search/srchengines.htm

3. How to Search the Web: http://daphne.palomar.edu/TGSEARCH

   A good guide to search tools.

4. NetLearn: www.rgu.ac.uk/~sim/research/netlearn/callist.htm

   A British site for resources for learning the Internet.

5. FindInfo: www.lib.berkeley.edu/Teaching.lib/Guides/Internet/FindInfo.html

   An intermediate level guide to finding information on the Internet.

6. Search Engine Watch: www.searchenginewatch.com

An advanced guide. Everything you ever wanted to know about search engines and related Internet research tools.

There are an estimated 1200 search engines on the Internet, and you can find 1173 of them listed and described on another search engine entitled SearchEnginesWorldwide. However, the below listed search engines, are probably the best ones to start with.

RECOMMENDED SEARCH ENGINES, DIRECTORIES, PARALLEL AND METASEARCH SITES

AltaVista	AOL Netfind	Ask Jeeves
ByteSearch	Canada.com	Debriefing
DefenseLink	DisInformation	Dogpile
Electronic Search	Euroferret	Excite
Go 2 Net	Go Network	Highway 61
HotBot	Huskeysearch	ICQiT
Infind	Infobase International	Infomak
Infoseek	iSleuth	Jughead
LookSmart	Lycos	Magellan
Mamma	Megacrawler	Metacrawler
Metafind	MSN	myGO
Netscape Search	Northern Light	Onesearch
OneSeek	Planet Search	Search.com
Sherlockhound	Snap	Studyweb
Veronica	Webcrawler	WebTV
Whatuseek	World Wide News	Yahoo
Z-search	800go	

The results of a search conducted through a search engine will provide a list of 'hits' in response. Typically, an item will be shown with a text title, followed by one to two lines taken from the web page, and lastly it will provide the web site address. The web site address is 'hypertext', which can be clicked on to go to the selected site. Many web sites also have 'hot links' which act as 'hypertext' and will take you further into specific related subjects.

"A Guide to Internet Web Sites of Counterintelligence Interest"
by John C. Morley, US Army detailee to NACIC, PIO, Community Training Branch,
(703) 874-4083. This replaces the previous edition dated 11 January 2000.